

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

3

(43)Date of publication of application : 26.02.1987

G06K 9/62

(71)Applicant : NEC CORP

(72)Inventor : KAMI HIROYUKI

(57)Abstract:

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Translated Excerpt of Citation 3

Japanese Patent Laid-Open Publication No. SHO 62-044897

From Page 6, Line 4 to Line 17:

“Action

According to the present invention, when a collation dictionary is represented by feature vectors, comparison is performed, in the middle of a distance calculation, between the distance obtained at the point of time and the maximum distance in the candidate category obtained by collation up to the present. If the distance at the point of time is larger than the maximum distance, the candidate category is detected without calculating the distance of the residual part. The distance is found by sum of products operation with the vector, and time taken by the comparison between two of these distances is far shorter than the time for the sum of products operations. Accordingly, in the present invention, if the distance up to the way of the distance calculation is larger than the maximum distance, the sum of products operation for the residual part is suspended, and then another distance calculation with the next vector is started, thereby the collating time is shortened.”

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭62-44897

⑮ Int.Cl.⁴

G 06 K 9/62

識別記号

庁内整理番号

B-6972-5B

⑬ 公開 昭和62年(1987)2月26日

審査請求 未請求 発明の数 1 (全5頁)

⑭ 発明の名称 文字認識装置

⑯ 特 願 昭60-184847

⑰ 出 願 昭60(1985)8月22日

⑱ 発 明 者 上 博 行 東京都港区芝5丁目33番1号 日本電気株式会社内
⑲ 出 願 人 日本電気株式会社 東京都港区芝5丁目33番1号
⑳ 代 理 人 弁理士 本庄 伸介

明 細 書

1. 発明の名称

文字認識装置

2. 特許請求の範囲

(1) 入力された未知パターンに類似する候補カテゴリ名を求める文字認識装置において、前記未知パターンから特徴を抽出し特徴ベクトルに変換する特徴抽出手段と、認識対象カテゴリの標準パターンの特徴を抽出し得られた特徴ベクトルとその標準パターンのカテゴリ名とを記憶する辞書記憶手段と、距離の小さい順に定められた数の距離及びカテゴリ名を記憶する候補記憶手段と、前記未知パターンの特徴ベクトルと前記辞書記憶手段に記憶されている一つの標準パターンの特徴ベクトルとの距離を求めるのにベクトルの次元を分け、分けた途中までの累積距離が前記候補記憶手段にある距離のうちの最大の距離より大きいときは次の前記標準パターンの特徴ベクトルとの距離

計算に制御を移し、前記累積距離が前記最大距離より小さいときは距離計算を続行し最後の次元までの距離と前記標準パターンのカテゴリ名とを出力する距離計算制御手段と、前記距離計算制御手段からの距離と前記候補記憶手段にある距離との比較により前記候補記憶手段にある距離とカテゴリ名との更新を行なう候補制御手段とを有することを特徴とする文字認識装置。

(2) 特許請求の範囲第1項記載の文字認識装置において、前記候補記憶手段の候補カテゴリ数を距離の小さい順に並列された2個とすることを特徴とする文字認識装置。

(3) 特許請求の範囲第1項記載の文字認識装置において、前記候補記憶手段の候補カテゴリ数を2個とし、前記候補制御手段は距離計算制御手段からの距離と、前記候補記憶手段の第1位の距離に両値を加えた値と第2位の距離とのうちの小さい方とを比較して、前記距離計算制御手段における前記距離計算を続行するか否かを制御することを特徴とする文字認識装置。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は、認識対象カテゴリ数が多い場合に適した文字認識装置に関する。

(従来技術とその問題点)

認識対象カテゴリ数が多くなると候補カテゴリ検出速度の低下が問題となる。従来候補カテゴリ検出速度の高速化のために次のような手法がとられてきた。

(a) 段階的な検出すなわち最初の段では少数の大局的な特徴によって変わされた照合辞書との照合で候補カテゴリを求め、次の段では次に大局的な特徴によって変わされた照合辞書との照合で候補カテゴリを絞り、次の段ではさらに微少な特徴にて変わされた照合辞書との照合で絞られた候補カテゴリの内て絞っていくのを繰り返す方法で、照合対象カテゴリ数を減らす。

(b) 照合辞書はカテゴリごとでなく類似している特徴をもつカテゴリはまとめて統合した照合辞書として仮のカテゴリ名を付け(a)と同様に

段階的に検出することで照合回数を減らす。

(c) 認識対象カテゴリの照合辞書を使用頻度に応じて分類し、頻度の多い分類内で類似の程度の高いカテゴリがあれば頻度の少ない分類にある照合辞書との照合を行なわないことで照合対象カテゴリ数を減らす。

(d) 照合辞書との距離を計算するのにオーバーフロー検出付きのレジスタを使用し距離計算の途中でオーバーフローが生じたら次の照合辞書との距離計算に制御を変更する。

(a)、(b)及び(c)は各カテゴリ又は仮のカテゴリの照合辞書の長さによって速度が決まってしまう、(d)はレジスタのビット長によって速度の向上が限定されるという問題点があった。

そこで、本発明の目的は、照合辞書との照合計算を照合辞書の長さによって分けて分けられたところまでの照合の程度で次の照合辞書との照合に制御し、候補カテゴリが高速に検出できる文字認識装置を提供することにある。

(問題点を解決するための手段)

-3-

前述の問題点を解決するために本発明が提供する手段は、入力された未知パターンに類似する候補カテゴリ名を求める文字認識装置であって、前記未知パターンから特徴を抽出し特徴ベクトルに変換する特徴抽出手段と、認識対象カテゴリの標準パターンの特徴を抽出し得られた特徴ベクトルとその標準パターンのカテゴリ名とを記憶する辞書記憶手段と、距離の小さい順に定められた数の距離及びカテゴリの名を記憶する候補記憶手段と、前記未知パターンの特徴ベクトルと前記辞書記憶手段に記憶されている一つの標準パターンの特徴ベクトルとの距離を求めるのにベクトルの次元を分け、分けた途中までの累積距離が前記候補記憶手段にある距離のうちの最大の距離より大きいときは次の前記標準パターンの特徴ベクトルとの距離計算に制御を移し、前記累積距離が前記最大距離より小さいときは距離計算を続行し最後の次元までの距離と前記標準パターンのカテゴリ名とを出力する距離計算制御手段と、前記距離計算制御手段からの距離と前記候補記憶手段にある距

-5-

-4-

離との比較により前記候補記憶手段にある距離とカテゴリ名との更新を行なう候補制御手段とを有することを特徴とする。

(作用)

本発明では、照合辞書が特徴ベクトルで表現されているときに距離計算の途中で、途中までの距離と今までの照合で得られた候補カテゴリの内での最大の距離との比較を行ない、途中までの距離の方が大きければ残りの部分の距離計算を行なわないで候補カテゴリを検出する。距離はベクトルの積和演算で求め、積和の途中までの距離と最大距離との比較を行なっても比較の時間は積和の時間に比べて無視できるほど小さい。そこで、本発明では、途中までの距離が最大距離より大きければ、残りの部分に対する積和演算を中止して次のベクトルの距離計算に入るようにすることで、照合時間の短縮を計っている。

候補カテゴリが求まると、候補カテゴリの距離の中から最小距離 d_1 と次最小距離 d_2 とを検出し、絶対閾値 T_1 と相対閾値 T_2 によって次の判定を行な

-6-

うのが一般的である。

① $d_1 \leq T_1$ かつ $d_1 - d_2 \geq T_1$ のときには最小距離 d_1 であるカテゴリ名を判定結果とする。

② $d_1 > T_1$ であれば読取り不能を判定結果とする。

③ $d_1 \leq T_1$ かつ $d_1 - d_2 < T_1$ であれば、 d_1 と d_2 の距離をもつカテゴリを次のペア判定手段に送るか、または読取り不能を判定結果とする。

従って類似パターンの数が最大2個に限定できる場合、候補カテゴリ数や相対閾値 T_2 を候補カテゴリ検出に使用することにより、本発明ではさらに高速化が計れる。

(実施例)

以下、本発明の実施例を図面を参照して詳細に説明する。

第1図は本発明の一実施例の装置のブロック図である。同図において信号10は帳票の反射光で、観測手段1へ入力される。観測手段1は帳票上のパターンを光学的に走査して信号10を得て、信号10を光電変換する。光電変換して得られる電気信

-7-

数の距離と対応する標準パターンのカテゴリ名とを記憶しておく記憶部をもち、複数個ある距離の記憶部は判定処理開始前に大きな値で初期化しておく。距離計算制御手段4は、信号21の前記文字パターンからの特徴ベクトルと信号31の辞書の特徴ベクトルとの距離計算を行なうのに、まず予め決めた回数要素間距離までの累積である距離を求める。次に求まった距離と候補記憶手段5からの最大の距離である信号51との比較を行なう。比較して最大の距離の方が小さければ、残りの要素に対する距離計算を中断して辞書記憶手段3から次の標準パターンの特徴ベクトルとの距離計算に制御を移す。また、最大の距離の方が大きければ、前記決められた回数の距離計算処理を繰り返す。そして、全部の次元までの距離計算が繰り返されたときは、得られた距離と信号31として入力した標準パターンのカテゴリ名とを信号41として出力する。

候補制御手段6は候補記憶手段5に記憶されている距離の小さい順に並べられた複数個の距離と

号について票面の白の部分は“0”に、黒の部分は“1”に量子化して2値信号を生成することにより、1文字ごとに切出した文字パターンを得る。この文字パターンを表わす2次元の配置にある“0”と“1”を信号11として特徴抽出手段2に出力する。特徴抽出手段2は、信号11からその未知の文字パターンの特徴の抽出を行って特徴ベクトルを得、この特徴ベクトルを信号21として距離計算制御手段4に出力する。辞書記憶手段3は、認識対象カテゴリの標準パターンから特徴抽出し得られた特徴ベクトルと対応するカテゴリ名とを記憶していて、前記文字パターンが特徴ベクトルに変換され距離計算制御手段4に出力された後で以下の検出処理が開始され、このとき記憶している特徴ベクトルとカテゴリ名とを一つづつ信号31として出力する。

2つの特徴ベクトル間の距離は、対応するベクトル要素間距離を特徴ベクトルの次元数だけ累積することによって求まる。

候補記憶手段5は、距離が小さい順に決まった

-8-

信号41の距離とを比較して第1位と第2位の距離とカテゴリ名の更新を行なう。

上記処理の流れ図を第2図(b)に示す。

また候補カテゴリ数が2個にでき、第1位の距離と第2位の距離の差が閾値以上でなければならぬという条件を使用するとさらに高速化が可能である。距離計算制御手段4、候補記憶手段5及び候補制御手段6は次の処理を行なう。

候補記憶手段5は距離が小さい順に第1位と第2位の距離と対応する標準パターンのカテゴリ名とを記憶しておく記憶部をもち、二つの距離の記憶部は判定処理開始前に大きな値で初期化しておく。距離計算制御手段4は信号21の文字パターンからの特徴ベクトルと信号31の辞書の特徴ベクトルとの距離計算を行なうのに、まず予め決めた回数の要素間距離までの累積である距離を求める。次に求まった距離と候補記憶手段5からの第1位の距離に閾値を加えた値と第2位の距離のうちの小さい側である信号51との比較を行う。前記信号51が小さければ残りの要素に対する距離計算

-9-

-677-

-10-

を中断して辞書記憶手段 3 からの標準パターンの特徴ベクトルとの距離計算に制御を移す。又、信号 51 が大きければ前述の決められた回数の距離計算処理を繰り返し、全部の次元までの距離計算が繰り返されたときは得られた距離と信号 31 とを入力した標準パターンのカテゴリ名とを信号 41 として出力する。

この場合の処理は、同様に流れ図で表わすと第 2 図 (c) となる。

4、5 及び 6 は実際にはマイクロプロセッサで構成されたシステムのプログラムによって処理される。第 3 図は第 1 図実施例の要部をマイクロプロセッサで構成してなるシステムの具体例のブロックである。

第 3 図において、401 はマイクロプログラムを格納する ROM とシーケンス制御手段で、402 は ALU (算術論理演算回路)、403 は距離とカテゴリ名を一時的に記憶する RAM であり、404 はデータベース 404 を通してデータの受け渡しを行なう。候補カテゴリ名はインタフェース 701 を

介して出力され、また特徴抽出手段 2 はインタフェース 201 を介してつながっていて、入力パターンの特徴ベクトルが入力される。

(発明の効果)

以上説明したように本発明によれば距離を累積する途中で距離が大きくなれば残りの次元についての距離計算が不要となるので認識時間の短縮が計れ効果は大である。また従来の距離計算の一部を変更するだけで済むので効率的である。本発明は複数回の距離計算で得られる累積距離との比較に限定されず一回ごとに比較しても良いことは言うまでもない。

上記実施例では文字パターンから特徴抽出を行ない得られる特徴ベクトルで距離計算を行なっているが、文字パターンの "0" と "1" とを用いて距離計算する方法も利用出来る。

また本発明は文字に限らず一般のパターンに対しても適用可能である。

4. 図面の簡単な説明

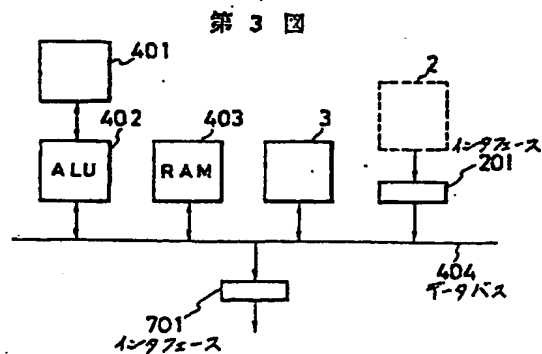
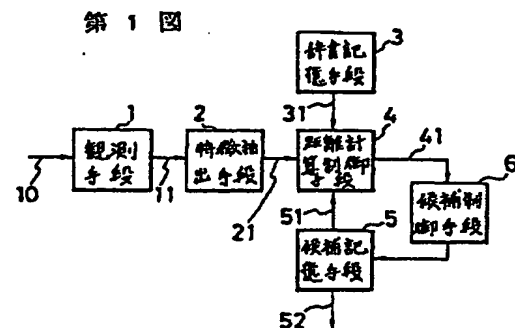
-11-

第 1 図は本発明の一実施例のブロック図である。第 2 図 (a) ~ (c) は第 1 図実施例の要部の処理を示す流れ図である。第 3 図は第 1 図実施例の要部を詳細に示すブロック図である。

図において、1 は観測手段、2 は特徴抽出手段、3 は辞書記憶手段、4 は距離計算制御手段、5 は候補記憶手段、6 は候補制御手段である。

代理人 弁理士 本 庄 伸 介

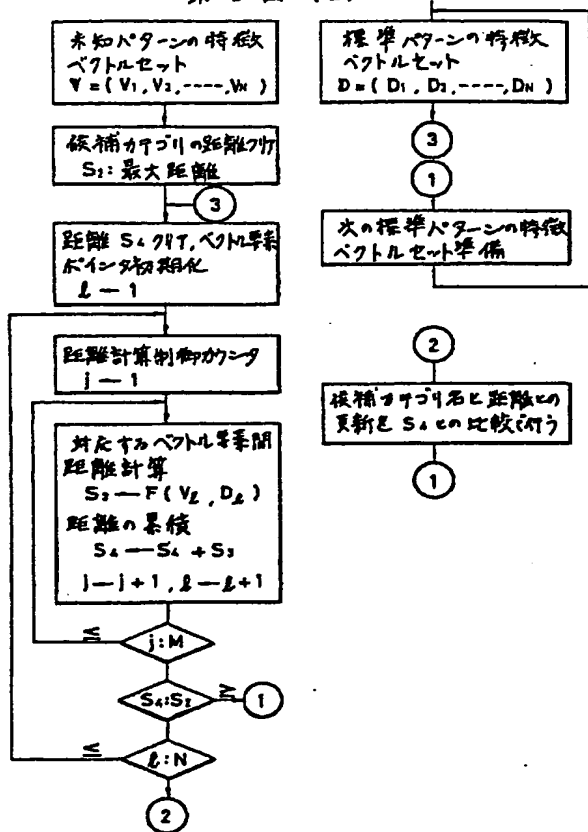
-12-



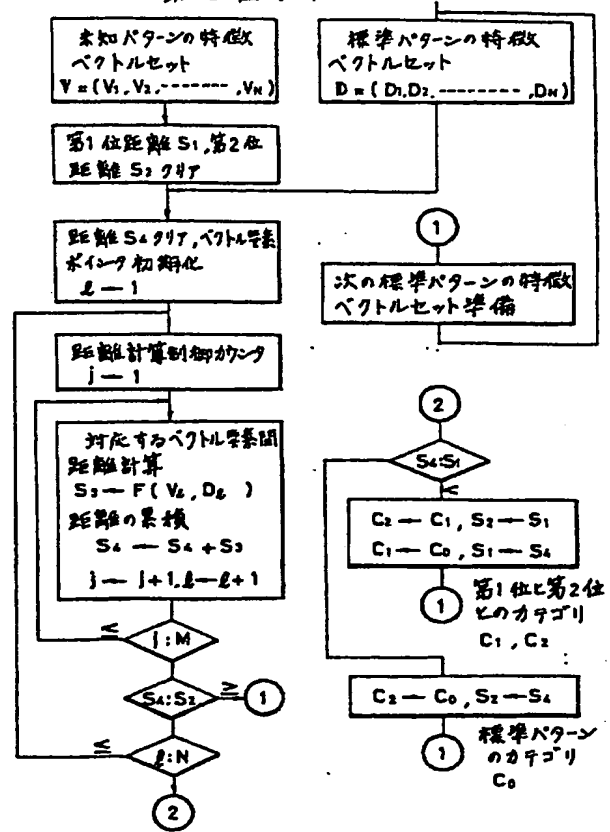
-13-

-678-

第2図 (a)



第2図 (b)



第2図 (c)

